TECHNICAL REPORT 1993
December 2010

Stress Response as a Function of Task Relevance

Steven A. Murray
Matthew Yanagi
Wayne Ensign
SSC Pacific

Burcu Darst
Gray Sourcing

SSC Pacific

SSC PACIFIC
San Diego, California 92152-5001

J. J. Beel, CAPT, USN
Commanding Officer

C. A. Keeney
Technical Director

ADMINISTRATIVE INFORMATION

This report was prepared for prepared for the Office of Naval Research, Arlington, VA, by the Applied Research Branch, Space and Naval Warfare Systems Center Pacific, San Diego, CA.

Released by
G. W. Anderson, Head
Applied Research Branch

Under authority of
M. J. Machniak, Head
Advanced Systems and Applied
Sciences Division

LH

TECHNICAL REPORT 1993
December 2010

Stress Response as a Function of Task Relevance

Steven A. Murray
Matthew Yanagi
Wayne Ensign
SSC Pacific

Burcu Darst
Gray Sourcing

Approved for public release; distribution is unlimited.

SSC Pacific
San Diego, CA 92152-5001

EXECUTIVE SUMMARY

A laboratory study was conducted to evaluate the impact of task relevance—the relation between task characteristics and an individual's skill set—on the stress response of Marine Corps Infantrymen, and to evaluate the sensitivity of several stress measures to distinctions of task relevance. The stress task involved a difficult simulated ground combat mission using Virtual Battlespace 2 (VBS2), presented on a laptop computer. Stress response measures included two self-report questionnaires and two salivary hormones (cortisol and nerve growth factor) collected at several points in the testing session.

A *high task relevance* group (Marine Infantrymen) and a *low task relevance* group (civilian volunteers) completed a set of VBS2 missions to generate elevated stress levels. Responses were compared to a cohort of civilians who were not given the VBS2 task (*control* group) to confirm the effectiveness of the stress manipulation, while the VBS2 groups were compared with each other to evaluate the basic hypothesis regarding task relevance. Marines were further partitioned into three sub-groups based on operational experience (i.e., combat experience, deployment experience but no combat, and no deployment experience) and compared with each other to evaluate the granularity of the task relevance hypothesis and to determine the sensitivity limits of the stress measures.

Results confirmed that the VBS2 task was effective in generating a significant stress response. Differences in stress response were also found between Marines and civilians, and between sub-groups of Marines, that supported the concept of task relevance. Each stress measure proved effective in some, but not all, comparisons, which highlighted the caution required in applying and interpreting such measures, or in relying on any single tool for general stress evaluation. Results are discussed in terms of stress models and the potential of stress measurement tools for military field use.

CONTENTS

Figures

Table

INTRODUCTION

Both research literature and military documentation contain abundant evidence regarding the effects of chronic and acute stress on operational decision making (Orosanu and Backer, 1996; U.S. Army, 1998). More specifically, stress contributes to decrements in learning, memory, and logical reasoning (e.g., Leach, 2004; Lieberman, et al., 2005). Because acute stress can be manipulated in the laboratory (e.g., Stokes, Kemper, and Marsh 1992; Hancock and Szalma, 2003), much can be learned from a systematic examination of this form of stress and the responses that it generates. Such research can lead to more sophisticated stress models, to improved measurement and evaluation tools, and to better understanding of how individual and organizational characteristics interact with stressors. In turn, this knowledge can diminish the impact of stress through more focused interventions (such as structured training events to manage individual stress response) resulting in more effective combat operations.

A useful perspective for both laboratory and field applications is the definition of stress as a mismatch between the demands of a task and the individual's ability to cope with those demands (Lazarus and Folkman, 1984)—the greater the mismatch, the greater the stress. Determining the mismatch, however, is an individual act. That is, each event is evaluated and compared to a *personal* assessment of one's abilities, which results in a unique stress response. Therefore, stress response is a function of personal judgment, or appraisal, and any approach to characterizing stress must account for the individual psychological factors that enter into that response. Furthermore, such factors can be very diverse; members of the same group may evaluate a common stressor event differently, and each will have a different stress response as a result (Driskell and Salas, 1991).

TASK RELEVANCE

These two factors (i.e., stress as a mismatch between demands and coping skills, and the role and variability of individual appraisal) come together as *task relevance*, a term defined for this study (but based on considerable anecdotal evidence). That is, if the significance of a stressor is a function of perceived mismatch between demands and capabilities, then the stress response should vary according to the degree that a task is personally perceived as related to one's capabilities. Individual appraisal of both demands and coping skills would differ, for example, when faced with a familiar task versus one that had never been encountered before, or would differ for those with confidence in their skills versus those who were uncertain of their mastery.

The definition of task relevance, as proposed above, involve two component features important to military performance:

- <u>Understanding</u> – An individual who grasps the significance and potential consequences of poor task performance may have a larger stress response than someone who does not perceive such consequences (e.g., Ritter, Riefers, and Schoelles, 2005). Understanding influences the accuracy of the task demand.
- <u>Skill</u> – An individual who is trained or otherwise more familiar with specific task demands may have a smaller stress response than someone who does not know how to perform that task, because their perceived mismatch, between demand and coping ability, is smaller; that is, the trained person has a greater sense of task control (e.g., Salas, Driskell, and Hughes, 1996). Skill influences the judgment of coping ability.

A professional pilot, for example, may better understand the real consequences of failure when operating a challenging flight simulation (e.g., "crashing") than an individual who knows little about the flying task. The pilot may therefore show a greater stress response than that of the naïve individual (for whom the simulation is only a game), based on task understanding. Because the pilot

possesses greater training and experience, however, the stress response to simulation demands may be smaller than that of someone without such advantages, based on personal skill. At a practical level, these factors must be combined to predict a likely overall response and, given human variability, may have to be combined at the individual or small group level to be of any use. Although the relative contributions of understanding and skill, and the nature of their interaction, are not known well enough to support response prediction, they can be tested through careful study of group or individual characteristics, as in the pilot and non-pilot example, above).

MEASURING THE STRESS RESPONSE

Stress response signals can be detected with both psychological and physiological tools. Because no single measurement tool is effective, or even appropriate, in all research circumstances, a variety of measures is often employed to fully characterize a stress response pattern.

Psychological states can be readily determined by asking an individual to state them. A primary method for such data collection is the self-report questionnaire, which requires an individual to rate his/her current state (e.g., feelings, attitudes, mood, etc.) via a standard checklist or set of short response items. Questionnaires are inexpensive, have good face validity (i.e., appear to measure what they purport to measure), and can be completed quickly—an important feature for gathering several measures in a short period of time. Data from questionnaires that are widely used (such as in research, clinical, or educational settings) can be accumulated over time to establish group and population norms, and can be benchmarked for validity and reliability. The State-Trait Anxiety Index (or STAI; Spielberger and Sydeman, 1994) is a popular self-report instrument of this type, and provides distinct scores for *state* anxiety (a property of the situation) and *trait* anxiety (a property of the individual) using a 4-point rating scale. The STAI requires a special scoring procedure before results can be interpreted, but has been effectively used for comparative studies of both anxiety and acute stress (e.g., Noto, et al., 2005; Chiffer McKay, et al., 2010). Other available instruments use a 5- or 10-point rating scale and can be evaluated without formal scoring (e.g., Kirschbaum, et al., 1995; Van Dongen, et al., 2004). The questionnaire described in Wang, et al. (2005), for example, contains 10-point scales for each of several stress-related dimensions—Stress, Anxiety, Effort, Frustration, and Difficulty—all on a single page. Disadvantages of self-report measures include misunderstanding or misinterpretation of questionnaire items and the possibility of deception or bias.

Physiological states reflect body reactions to psychological or physical stimulation. Because physiological processes are generated internally (i.e., through neural or biochemical mechanisms), their measurement bypasses the consciously mediated responses required by self-report questionnaires. Physiological methods can therefore be used in human research as a substitute for, or complement to, psychological approaches. Typical performance measurement methods for physiology include cardiac function (such as heart rate and blood pressure; e.g., Vrijkotte, van Doornen, and de Geus, 2000) and analysis of blood or salivary hormones (such as cortisol or human nerve growth factor; e.g., Kirschbaum and Hellhammer, 1999; Steptoe, Hamer, and Chida, 2007; Aloe, Alleva, and Fiore, 2002). Of these, salivary sampling is particularly attractive as a research measure because saliva is relatively easy to collect and requires almost no equipment. The primary disadvantage of physiological measures is that the body processes upon which they are based are influenced by many factors besides the stimulus of interest; dietary and drug habits, physical activity (even talking), state of health, time of day, etc. can dramatically alter physiological indices. Interpretation of physiological measures is also complex, as different mechanisms control different processes. Cortisol levels, for example, are controlled primarily by the hypothalamic-pituitary-adrenal (HPA) axis while nerve growth factor (NGF) levels are controlled primarily by the amygdala-medullary axis that, in turn, modulates the HPA (e.g., Aloe, Bohm, and Levi-Montalcini, 1986).

INITIAL RESEARCH

Our previous work (Murray, Ensign, and Yanagi, 2010) has confirmed the effectiveness of using a tabletop combat simulation—Virtual Battlespace 2 (VBS2; Bohemia Interactive, 2010)—to induce acute stress in a general population of civilians. After a brief training session, a pair of short, but extremely challenging, mission scenarios was presented to experiment participants. Stress response measurements were made at the beginning of the test session, after training, after each scenario, and after a short recovery period to establish stress profiles for each participant. This work yielded consistent stress response patterns for essentially all measurement methods and provided a foundation for more detailed investigations into specific shaping factors of stress, such as the task relevance.

The most direct path toward this goal would be to compare stress responses of two populations that differed only in the task relevance of a common stressor. In keeping with the earlier example regarding flight simulation and pilots, we propose that Marines should demonstrate a different stress response than civilians to a stress task that involves a military skill set, such as a ground combat mission. We reasoned that the VBS2 combat simulation would elicit a different stress response from a group of Marine Corps Infantrymen, a group of civilians who also completed the VBS2 task, and a group of civilians who did not receive any stressor at all (i.e., a control group). Furthermore, because understanding of consequences presumably changes with experience, we reasoned that Marines with operational combat experience should demonstrate a different stress response than Marines without such experience, owing to different levels of task understanding. The current study examines these hypotheses in terms of the following objectives:

1. **Select the most effective stress measures** to apply to the data set. While a comprehensive approach is desired, redundant or insensitive measures should be eliminated to retain focus on the analyses that follow.
2. **Calibrate the stress task** by applying VBS2 to new participant groups. Does VBS2 elicit a stress response in new cohorts? A significant result is essential before beginning any subsequent analyses.
3. **Evaluate the Marine stress response** by comparing Marine results with those of a civilian control group and with those of a civilian stress group**.** Do Marines demonstrate a different stress response than civilians exposed to the same combat simulation task? A significant result would provide a basic confirmation of task relevance as a factor in stress response.
4. **Evaluate the stress response at increasing levels of precision**. That is, how specific is the task relevance factor in stress response? Does combat or operational deployment experience yield a different pattern of stress responses, compared to Marines who have never been deployed? A significant result would further resolve the basic findings of Objective III.
5. **Evaluate differences in patterns of stress recovery**. This objective is exploratory only, and is included to better complete a picture of Marine Corps response to laboratory-induced stress. That is, if the pattern of stress response is different for Marines, is the pattern of recovery to pre-stress baseline levels also different?

METHOD

A factorial design was employed for the study to evaluate the basic effects of stress as a function of group characteristics. The stress task consisted of a simulated combat scenario, generated with the Virtual Battlespace 2 (VBS2) software currently employed for tactical training by the U.S. Marine Corps and other military services.

DESIGN

The factors of the experiment included:

- A two-level stress factor involving participants who completed the stress task, and other participants who did not (i.e., a civilian Control group).
- A two-level task relevance factor involving active duty Marine Infantrymen, trained in the skill set relevant to the stress task, and civilians who had no such training (i.e., a civilian Stress group).
- A three-level experience factor for the Marine group, involving those who had completed at least one combat tour (Combat sub-group), those who had been deployed but who had not received a combat rating (Deployed sub-group), and those who had never been operationally deployed (Not Deployed group).

The general design is shown in Figure 1.

<table>
<tr><th colspan="4">Stress</th><th rowspan="3">Control</th></tr>
<tr><th colspan="3">Marine</th><th rowspan="2">Civilian</th></tr>
<tr><td>Combat</td><td>Deployed</td><td>Not Deployed</td></tr>
</table>

Figure 1. Experiment Design.

Individual stress response was measured with a series of data collection events scheduled over the duration of the test session. Stress response measures (described earlier) included:

- The 20-item version of the STAI
- A multi-factor Stress Scale (as described in Wang, et al., 2005)
- Salivary components, gathered with a sublingual lozenge and salivette, and assayed for cortisol and NGF

PERSONNEL

The Commanding General, 1st Marine Division, authorized the use of active-duty Marine Infantrymen for the study, divided by the unit supervisor into combat and non-combat personnel. Civilian Stress group participants were recruited though an open news advertisement. Control group data were taken from a previous experiment employing an identical procedure; that is, new participants were not tested for this condition, but were drawn from an earlier research study involving an equivalent procedure.

Requirements for participants were based primarily on the need to control for external influences on alertness and diurnal hormonal cycles, and included:

- Males
- Age 18–30

- Sufficient rest in the previous 24 hours
- No medications (evaluated on a case-by-case basis)
- Non-tobacco users (or, at least no use within 24 hours)
- No caffeine on the testing day (evaluated on a case-by-case basis)

TESTING ENVIRONMENT

Participants were scheduled in groups of approximately five individuals, but tested individually at workstations in a common area. Each participant was supervised by an experimenter to ensure proper procedure execution and data collection, and to provide individual support as needed. Each experiment station included the stress task computer, instruction placards, salivette collection tubes, headphones, and an electronic timer, as shown in Figure 2.

Figure 2. Individual Experiment Station Layout.

PROCEDURE

All civilian participants were first screened with a telephone survey to ensure that basic participation requirements (above) were met prior to being admitted to the experiment. While participation requirements were also briefed to Marine Corps participants, compliance was not under the control of the experimenters and, therefore, was not guaranteed. All participants provided by the Marine Corps sponsors were accepted to the study and, like all civilian participants, were queried about their compliance with requirements prior to commencement of the experiment protocol.

Following check-in, completion of Informed Consent procedures, and orientation to the experiment environment, participants completed the following sequence of steps:

1. The experimenter asked each individual about their current state of general health and recent sleep status.
2. A (*Baseline*) saliva sample was collected, using commercial (Salimetrics, LLC) salivette tubes and standard collection procedures. The collection task required the participant to soak a small synthetic fiber lozenge under their tongue for 90 seconds, and then to spit it into a test tube. The experimenter provided instructions prior to sample collection, and ensured the proper soaking period with an electronic timer. Immediately following collection, all salivettes were stored in a freezer.

3. The participant was given instruction and experience with control procedures for the VBS2 task scenario, which represented a street environment in an Iraqi city. Control actions for the simulation were confined to simple horizontal movement in all directions, and to weapon (rifle) firing. Training was standardized by requiring the experimenter to read and demonstrate from a written script, and involved 5 minutes of hands-on execution by the participant as elements in the scenario were pointed out. All questions from the participant were answered during and following training. The participant wore sound-suppressing headphones, to reduce distractions, during all VBS2 activities.
4. A second (*Training*) saliva sample was collected, for later comparison with the *Baseline* sample, to identify any elevation of hormones following task training
5. The participant completed both STAI and Stress Scale forms to establish a pre-stress baseline. Because this collection occurred after the VBS2 training event, however, these data were also labeled as *Training* points for later evaluation.
6. The participant then completed the first of two actual VBS2 task scenarios. The mission required the participant to navigate a street in an Iraqi city to reach a safe destination amid potential insurgents and other threats. The participant was allowed to fire his weapon to protect himself, but was instructed to avoid confrontation, if possible. The experimenter warned that exactly six minutes was allowed to complete the mission. The experimenter further warned that the scenario would reset to the starting position each time the participant was "killed" in the simulation, but that the timer would <u>not</u> reset. The result of these rules was that the time pressure to complete the mission objective grew shorter with every "kill" event. No experimenter interaction or communication occurred during the VBS2 mission itself.
7. The participant completed a second Stress Scale form (*Stress 1*), to compare with the pre-stress baseline administration of step 5.
8. The VBS2 scenario was administered a second time, with the same six-minute limit. While all mission aspects were identical to those of the first session, participant navigation decisions and random elements of the VBS2 program itself ensured some variability in event flow. This step completed the stress phase of the experiment, and headphones were removed at this time.
9. Another (*Stress 2*) saliva sample collection, STAI, and Stress Scale were completed immediately following the second VBS2 mission, when stress was presumably at its highest level.
10. At this point, the participant was debriefed by the experimenter using a standardized written script regarding the extraordinary difficulty of the VBS2 mission and the complete purpose (i.e., stress manipulation) of the experiment.
11. Following a 15-minute delay period to allow the participant to return to a resting state, a final (*Recovery*) saliva collection, STAI, and Stress Scale were completed.
12. After the final sample collection, the experimenters answered any remaining questions, provided payment (civilians only), and released the participant.

A timeline of the experiment procedure—which lasted approximately 1 hour—is shown in Figure 3.

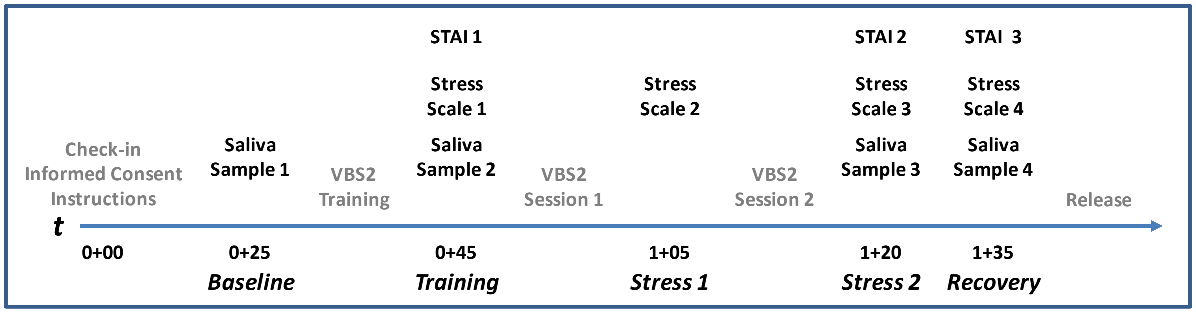

Figure 3. Experiment Procedure Timeline.

ANALYSIS

STAI and Stress Scale forms were scored by hand, using standardized rating procedures. Salivette tubes were shipped in dry ice to Salimetrics, LLC for specialized analysis. Each salivary sample was assayed in duplicate for cortisol and in triplicate for NGF, with results delivered to the experimenters as a spreadsheet data table. Mean values for all physiological data were transformed for analysis by converting to natural logarithm values (see Box, Hunter, and Hunter (1978).

A mixed-model, one-way analysis of variance (ANOVA) process was applied to the results to determine significant differences between groups, using both psychological and physiological measures described earlier. The three fundamental participant groups—Marine, Civilian Stress, and Civilian Control—were first examined, followed by systematic comparisons of each pair of groups. A similar strategy was applied to sub-groups of the Marine cohort—Combat, Deployed, and Not Deployed—to evaluate overall relationships and then pair-wise relationships. Although a significance level of $p = .05$ or better was initially selected as a default threshold for analysis, it is advisable in research practice to provide a higher threshold when testing multiple factors simultaneously, as many measurements increase the likelihood of finding a significant result by chance. A Bonferroni correction (see Bland, 1995) was therefore applied to the data, a procedure that involves dividing the nominal ($p = .05$) level of significance by the number of simultaneous tests being evaluated. The greater the number of tests, therefore, the higher the threshold (i.e., the lower the p value) required for a significant result.

RESULTS

The study was designed to include 40 Infantry-trained Marines (20 with operational combat experience and 20 without combat experience), 20 civilians who completed the VBS2 stress task (Civilian Stress), and 20 civilians who did not complete the task (Civilian Control).

A total of 37 of 40 scheduled Marines completed the study protocol, of which 36 provided usable data. Of this cohort, 20 were combat experienced (Combat sub-group), 6 had been deployed but had not received a combat rating (Deployed sub-group) and 10 were Infantry-trained but had never been deployed (Not Deployed sub-group).

Finally, 18 of 20 scheduled participants provided suitable data for the Civilian Stress group, and 21 data sets were included from a previous study to establish a Civilian Control group.

OBJECTIVE I – SELECTING STRESS RESPONSE MEASURES

The first step in the analysis involved an assessment of the performance of the stress response measures themselves. A Spearman correlation matrix for all measures is presented in Table 1. These correlations were calculated for all participant groups and all data, including the stress *Recovery* data point. Given eight comparisons in the correlation, the Bonferroni correction established a statistical threshold of p = .00625 (.05/8) for significance. Significant correlations are rendered in bold print.

Table 1. Spearman Correlation Matrix – All Data.

	STAI	**STRESS**	**ANXIETY**	**EFFORT**	**FRUSTRATION**	**DIFFICULTY**	**Cortisol**	**NGF**
STAI		**0.774**	**0.741**	0.166	**0.628**	**0.616**	0.052	-0.006
STRESS	**0.774**		**0.855**	**0.277**	**0.691**	**0.691**	0.006	0.055
ANXIETY	**0.741**	**0.855**		**0.233**	**0.674**	**0.641**	0.005	0.103
EFFORT	0.166	**0.277**	**0.233**		**0.284**	**0.323**	**0.245**	0.000
FRUSTRATION	**0.628**	**0.691**	**0.674**	**0.284**		**0.769**	0.090	-0.007
DIFFICULTY	**0.616**	**0.691**	**0.641**	**0.323**	**0.769**		0.046	0.001
Cortisol	0.052	0.006	0.005	**0.245**	0.090	0.046		0.160
NGF	-0.006	0.055	0.103	0.001	-0.007	0.001	0.160	

Multiple significant correlations indicate that many measures appear to be addressing similar underlying factors in self-reports regarding stress response. The highest correlations exist among the components of the Stress Scale and between these components and the STAI. The time profiles of each of the Stress Scale measures, taken from the entire data set over the duration of the experiment, are shown in Figure 4 and illustrate these similarities (with the notable exception of EFFORT).

Figure 4. Time Profiles: All Components of the Stress Scale.

The relatively low correlations of the EFFORT scale with other factors may reflect a ceiling effect, i.e., every participant was likely to state that they were exerting their maximum effort, which would maximize the scores and reduce the variability for this measure across the experiment. The EFFORT scale was also the only self-report measure that correlated significantly with a physiology measure (Cortisol), however, which may indicate that this component of the Stress Scale is sensitive to different factors than those of the other measures.

Physiology measures also showed low correlations with other factors. To the extent that physiology measures demonstrate significant, interpretable patterns in subsequent analyses, this result may indicate that physiology is capturing a different dimension of stress response than self-report tools.

As a result of these high correlations, only two self-report (psychological) measures—the STAI and the STRESS scale—were selected for use; although highly correlated with each other, the STAI and the STRESS scale are nevertheless independently developed tools whose separate and combined performance as stress measures are of practical interest.

Together with the two physiological measures, a total of four tools are therefore applied to the experiment results. A Bonferroni correction for test with these measures (.05/4) yielded a significance threshold of p = .0125 (i.e., much better than the p = .00625 if the entire measurement set were retained). Given the exploratory nature of this work, however, and the limited size of the testing population, we chose to present all results that met the basic standard of p = .05 or better, with the understanding that further work—involving more participants and greater experimental power—will eventually resolve the true value of the results reported here. *Results presented with p > .0125, however, should be evaluated with caution.*

Summary

Correlation analysis is a simple method for investigating relationships among multiple measurements and evaluating their utility. A correlation matrix of the results provided by the multiple measurement tools of this experiment showed that many scales could be deleted without losing analytical insights. Low correlations, however, may indicate potentially unique contributions of physiological measures to overall stress analysis. Four measurement tools were selected for the data analysis, and a statistical correction for simultaneous tests was deferred in the interests of identifying potentially valuable—if possibly weak—phenomena.

OBJECTIVE II – CALIBRATING THE STRESS TASK

The next issue of data analysis was to confirm that the VBS2 task was effective in eliciting a stress response among both Marine and civilian groups. The time patterns for selected response measures, diagrammed in Figure 5, show clear distinctions between the two stress groups and the control group, with *increasing* levels of reported stress for the psychological instruments (STAI and STRESS scale). As noted earlier, only the STRESS scale result is shown to represent the multi-factor questionnaire, although other components of the Stress Scale questionnaire showed the same profile and statistical significance. This convention is followed throughout the report to reduce length and to improve focus.

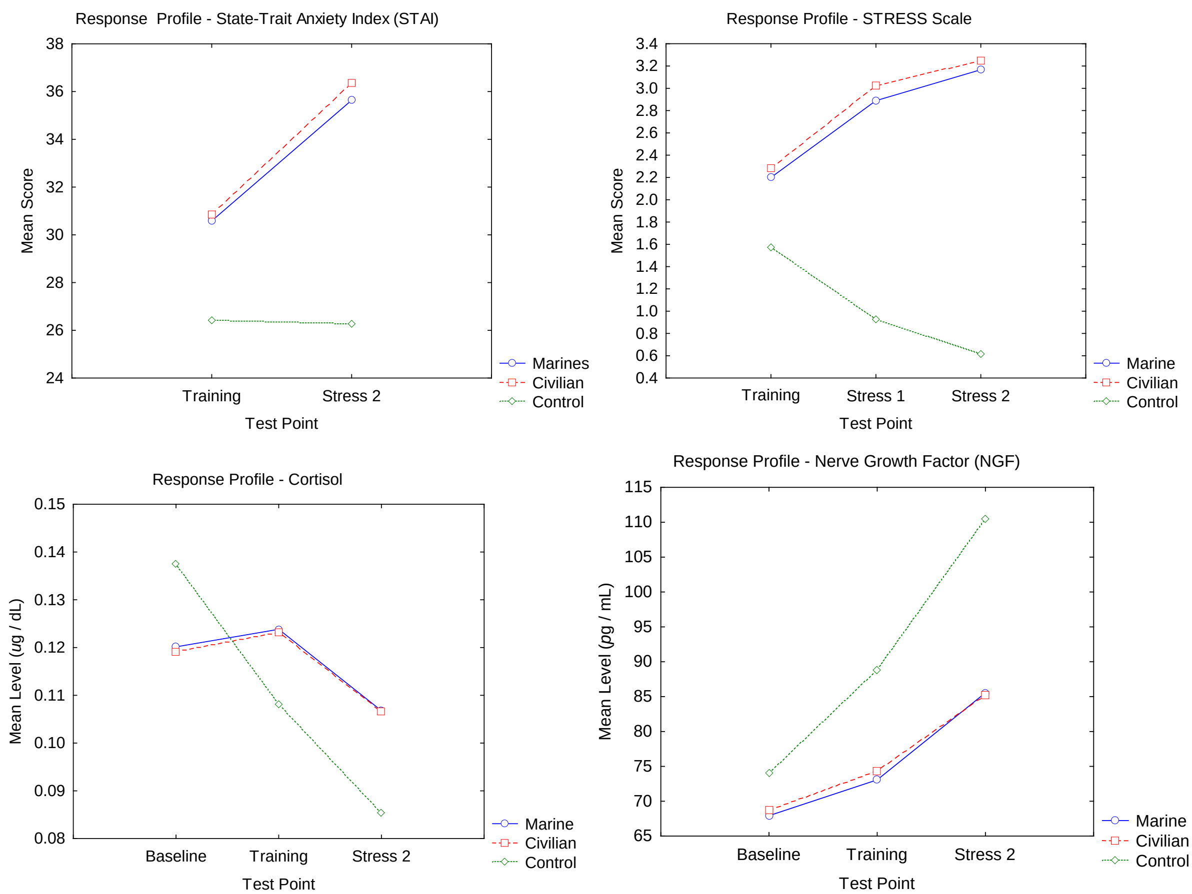

Figure 5. Time Profiles: Stress Response Measures Selected for Analysis.

The interpretation of the physiology measures is more complex. A distinction is seen between stress groups and the control group for these measures, which demonstrates sensitivity to the stress task, but cortisol shows *decreasing* levels across much of the testing session that was not aligned with expectations. The steady decline in Control group cortisol may indicate that physical activities prior to the experiment session might have elevated the cortisol in all participants, more than any testing events administered during the session, allowing levels to fall over time as participants sat at their workstations. Furthermore, although the initial VBS2 training may have elicited an initial cortisol elevation among Stress groups, the task challenge was not enough to activate HPA response any further, allowing the negative trend to manifest in the Stress groups as well. Without further analysis, however, this explanation is only conjectural.

While the patterns of nerve growth factor (NGF) show continuous elevation over the testing session, in accordance previous research, the uniformly higher levels for the Control condition were unexpected. Because the Control data were drawn from an earlier experiment, however, the generally higher levels of NGF might be explainable from differences in test conditions for the two experiments; because the original study was conducted in a windowless laboratory, and the earlier participants also wore a heart monitor apparatus, it is possible that overall stress levels may have been higher than for conditions of the study reported here.

General Effects

The effectiveness of the VBS2 as a stressor was evaluated statistically by comparing response data for the three experiment groups—Marines, Civilian Stress, and Civilian Control. A summary of significant one-way ANOVA tests for this three-group comparison is shown in Figure 6.

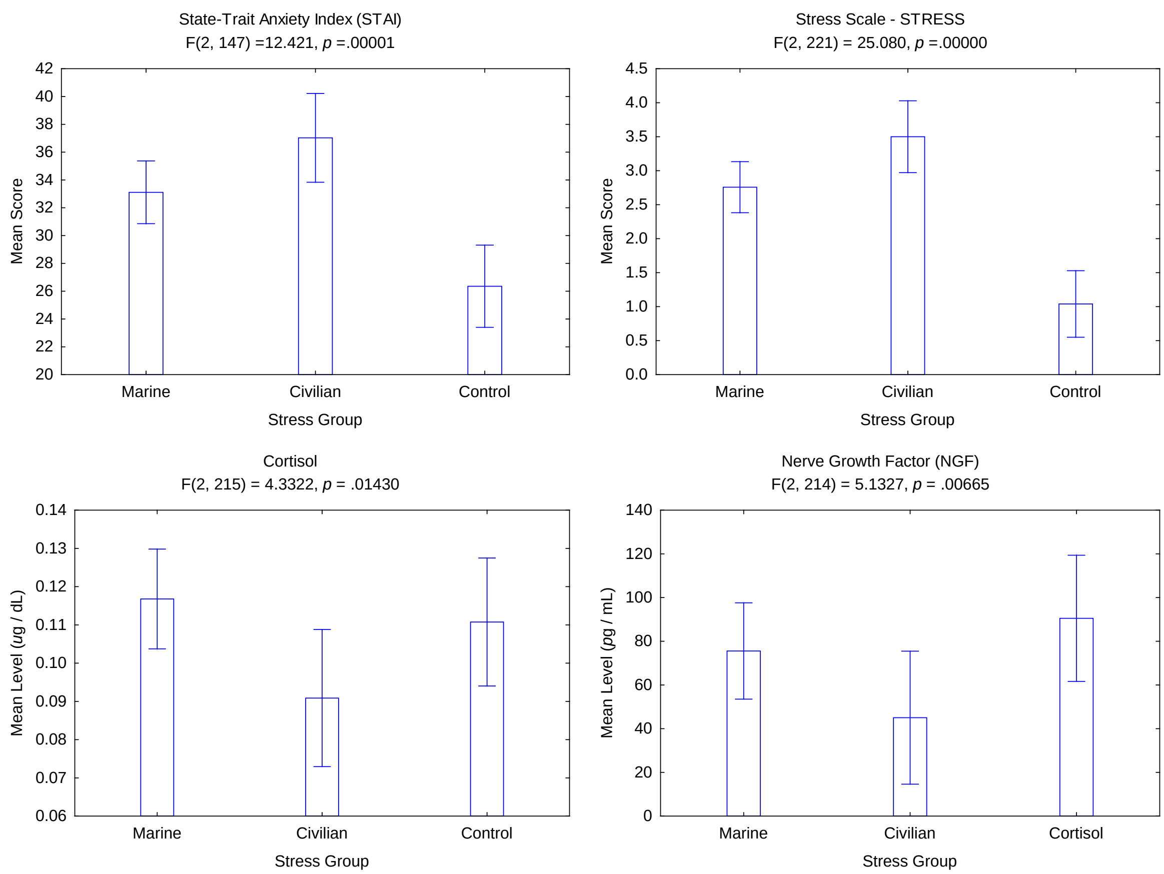

Figure 6. ANOVA Results: Group Comparisons – All Groups

Both Marines and the Civilian Stress groups showed elevated stress responses based on the self-report questionnaires. Physiological results were equivocal, however, as the greatest cortisol elevation was observed in the Marines but the greatest elevation in NGF was found in the Control

group. More detailed examinations are presented below to resolve these results, using pair-wise group evaluations.

Confirming the Stress Task

The most direct way to confirm that the VBS2 task was effective in eliciting a stress response is to compare results for equivalent groups. In this case, comparison is made between the Civilian Stress group and the Civilian Control group—a homogenous community that differed only in whether or not the VBS2 task was completed. One-way ANOVA results for this evaluation are shown in Figure 7.

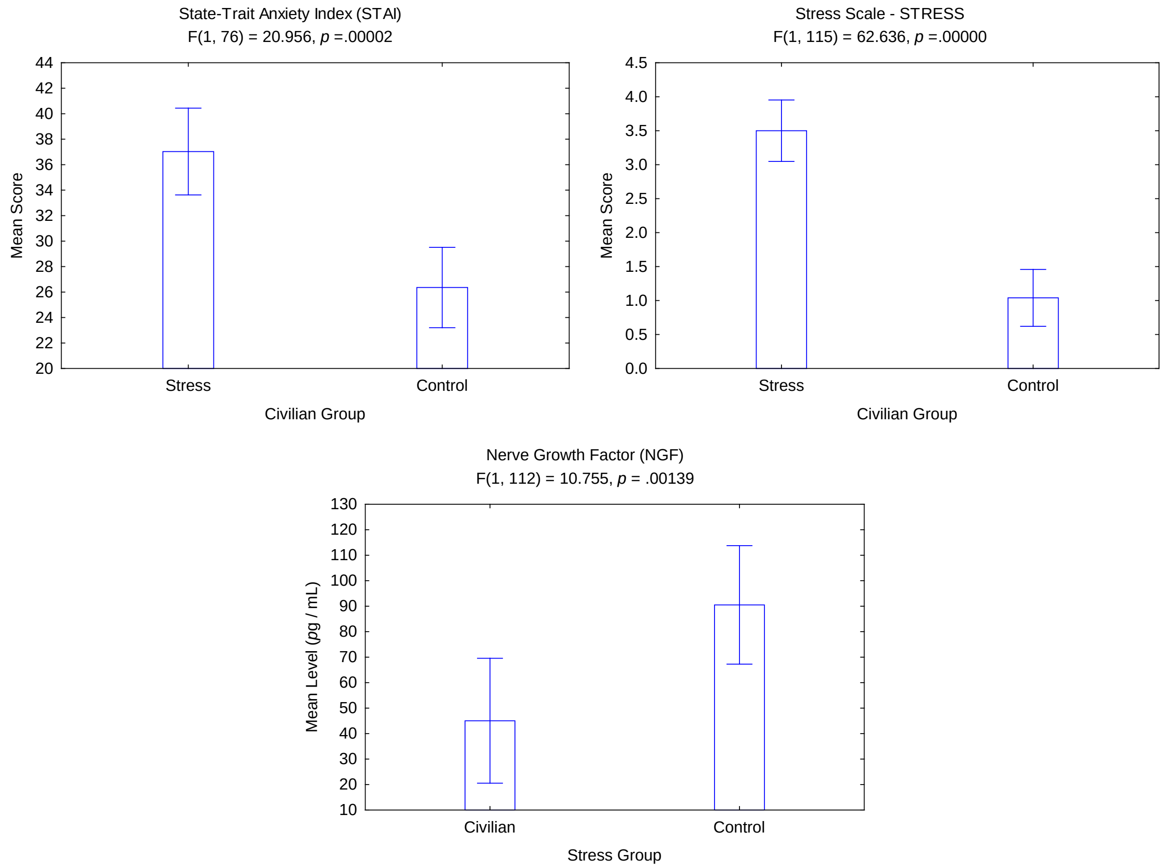

Figure 7. ANOVA Results: Group Comparisons – Civilian Stress and Civilian Control.

These ANOVA tests showed that both self-report measures distinguished between the Civilian Stress and Control groups, providing psychological evidence that the VBS2 task was effective as a stressor. Results for cortisol were not significant (i.e., even at the $p = .05$ level), however, and the direction of the NGF difference (i.e., higher levels for the Control group, as introduced in Figure 6), while significant, were anomalous. Although several explanations for these effects may exist (e.g., nature and intensity of the VBS2 task or environmental conditions of the experiment), such results, obtained in a most basic test of stress response, highlight the complexity of using physiology tools as performance measures without qualification.

Summary

These results provided general evidence that the VBS2 task was effective in eliciting a measurable response, using a variety of methods. While physiological measures were sensitive to stress, however, their sensitivity and interpretability preclude a consistent endorsement of their utility, at least within this experiment design.

OBJECTIVE III – EVALUATING THE MARINE STRESS RESPONSE

As shown in Figure 6, the Marine response to the stress task differed from that of the two civilian groups, and pair-wise group tests can help to more precisely characterize these differences.

Basic Marine Stress Response

Marine data are first compared with those of the Civilian Control group to establish the significance of the VBS2 as an effective stressor for the Marine community. One-way ANOVA results for this comparison are shown in Figure 8.

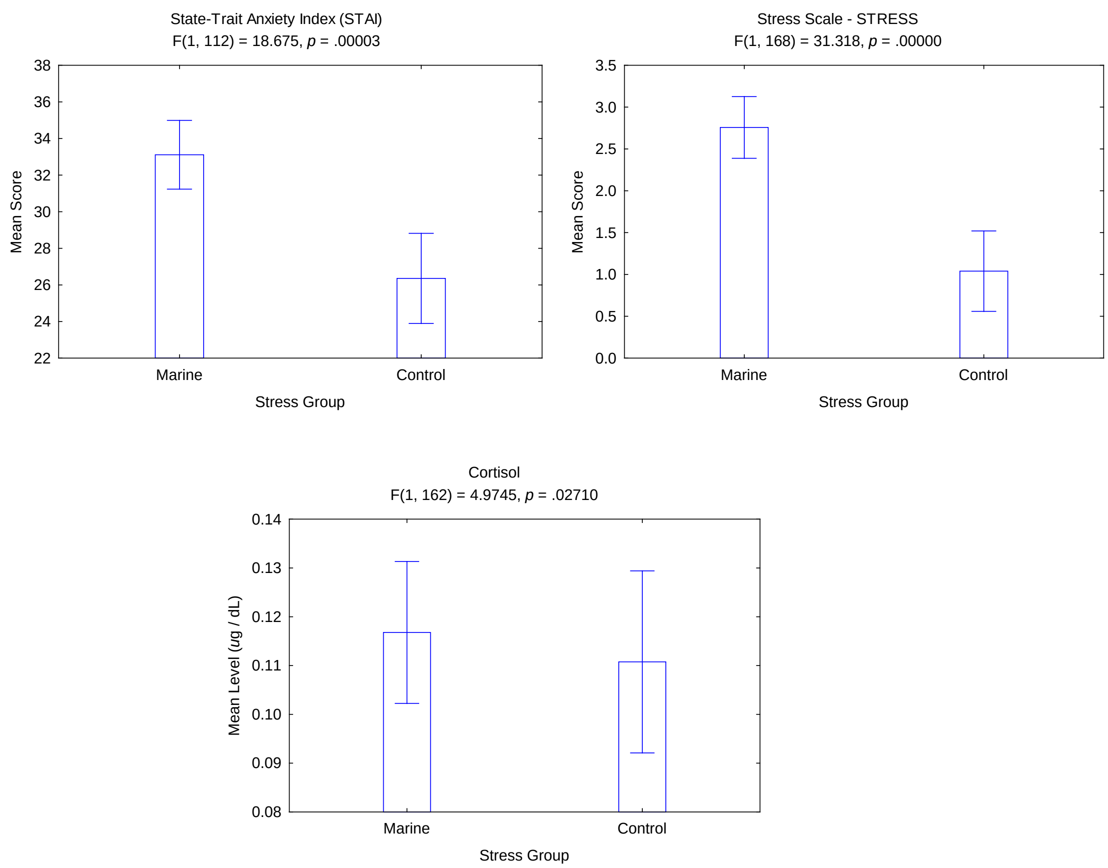

Figure 8. ANOVA Results: Group Comparisons – Marine and Civilian Control.

The self-report questionnaires for this test are in agreement with those of the earlier, general comparison (Figure 7), and show significantly elevated stress in the Marine group. Of the physiology measures, however, it was cortisol that demonstrated a significant effect index (i.e., showing a significant elevation in response to stress) and NGF that did not. The lack of NGF significance is, however, accounted for by the observed (and, as yet, unexplained) high NGF levels of the Control

group, shown earlier in Figure 5, which overwhelmed any VBS2 stress response of the Marine cohort.

Specific Marine Stress Response

A more challenging—but more interesting—distinction can be made by comparing the Marine group with the Civilian Stress group, yielding information about the *community-specific* nature of the stress response. Results of this test are summarized in Figure 9. While both physiological measures were significant in this comparison, only one self-report instrument—the STRESS scale—reached this threshold.

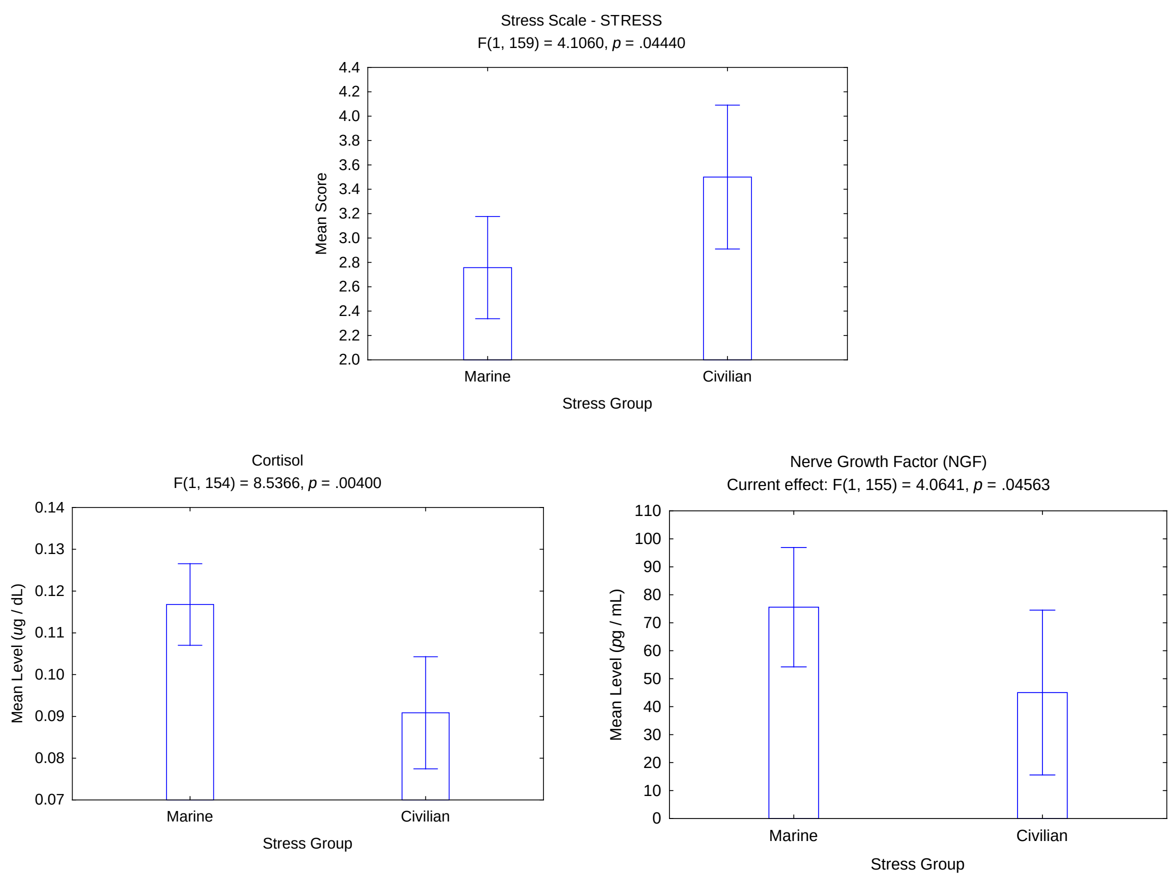

Figure 9. ANOVA Results: Group Comparisons – Marine and Civilian Stress.

As shown in these figures, the STRESS scale demonstrated a higher stress response in the Civilian Stress group than in the Marines, while both physiology measures showed the opposite effect (although changes in both cortisol and NGF were in the direction expected by previous research). These results may indicate a distinction between psychological ratings of the stress experience—which may involve personal decisions about how such an experience is reported to an experimenter—and physical responses to the same stimuli. It is worth noting the greater precision in the physiology results when comparing two groups that had both experienced the stressor task, compared to the earlier results that compared stressed and non-stressed groups; the profile of physiological substrates (i.e., cortisol and NGF) expressed in response to a series of stressor stimuli

may be influenced by different mechanisms than those expressed by simply sitting for a period of time.

Summary

These results provide evidence that the VBS2 was effective in eliciting a stress response among Marines and civilians, when compared with controls, as multiple measures indicated significant changes following task exposure. Interpretation of these results was not simple, however, as it appears that self-report measures may be effective in distinguishing between stressed groups and a control group, but only somewhat effective in detecting smaller differences within the (Marine and civilian) stressed groups. The self-report results further indicate that the psychological stress response was modulated among Marines, when compared with civilians, which bears on the central issue of task relevance in this experiment.

Although the patterns of physiological results were also difficult to interpret, as measures were inconsistent across group comparisons, demonstrated effects that were counter to other research literature (e.g., NGF, Figure 8), and contradicted a self-report measure (Figure 9). While NGF results for comparisons involving the Control group could, at least, be explained by the consistently high levels of this hormone in controls, attempts to integrate psychological and physiological measures for the VBS2 stress response are, for the moment, problematic.

OBJECTIVE IV – EVALUATING THE IMPACT OF OPERATIONAL EXPERIENCE

The analysis of the factors underlying stress response is extended here to include differences within sub-groups of the Marine participants. Specifically, comparisons among Combat, Deployed, and Not Deployed sub-groups of the Marine cohort are made using the same analysis process applied in the previous section. (Note that this approach treats the Not Deployed Marines as a functional control group, i.e., representing the least task-specific exposure to operational combat stress among Marine participants.)

General Effects

Results of a general evaluation of all three Marine sub-groups are shown in Figure 10.

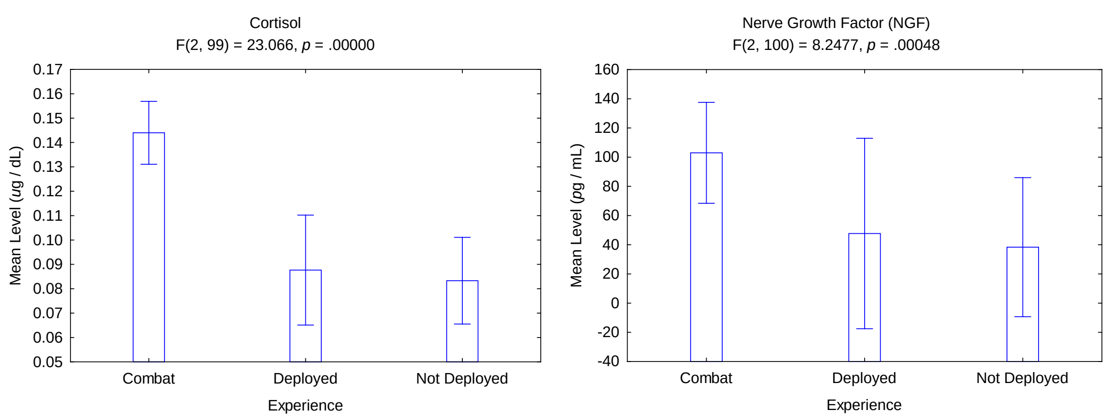

Figure 10. ANOVA Results: Sub-group Comparisons – All Marines.

Of the measures used for this analysis, only physiology results reached statistical significance. These measures appeared to reflect consistent and significant differences among Marine sub-groups, reflecting higher stress response levels with increasing levels of operational experience. Both cortisol and NGF levels were elevated in this analysis which, according to previous research, is indicative of higher stress response with increasing operational experience. Note that this test, like the comparison of Figure 9, involved only participants who completed the stress task (i.e., no control group was included) and also yielded more interpretable results.

Specific Experience Effects

The following analyses provide more detailed examinations of stress response differences through pair-wise comparisons among the three Marine sub-groups. These results are only exploratory, as the ANOVA tests discussed in "General Effects" above would typically not warrant more detailed examination of these sub-groups.

Combat versus No Deployment

The first analysis compared Marines with true combat experience (Combat) with Infantry-trained Marines who had never been operationally deployed (Not Deployed). ANOVA results for this test are shown in Figure 11. While the STAI self-report questionnaire was statistically significant, the STRESS scale was not. In fact, the STAI result did not reach significance to the Bonferroni threshold of $p = .0125$, indicating another weak result for self-report measures.

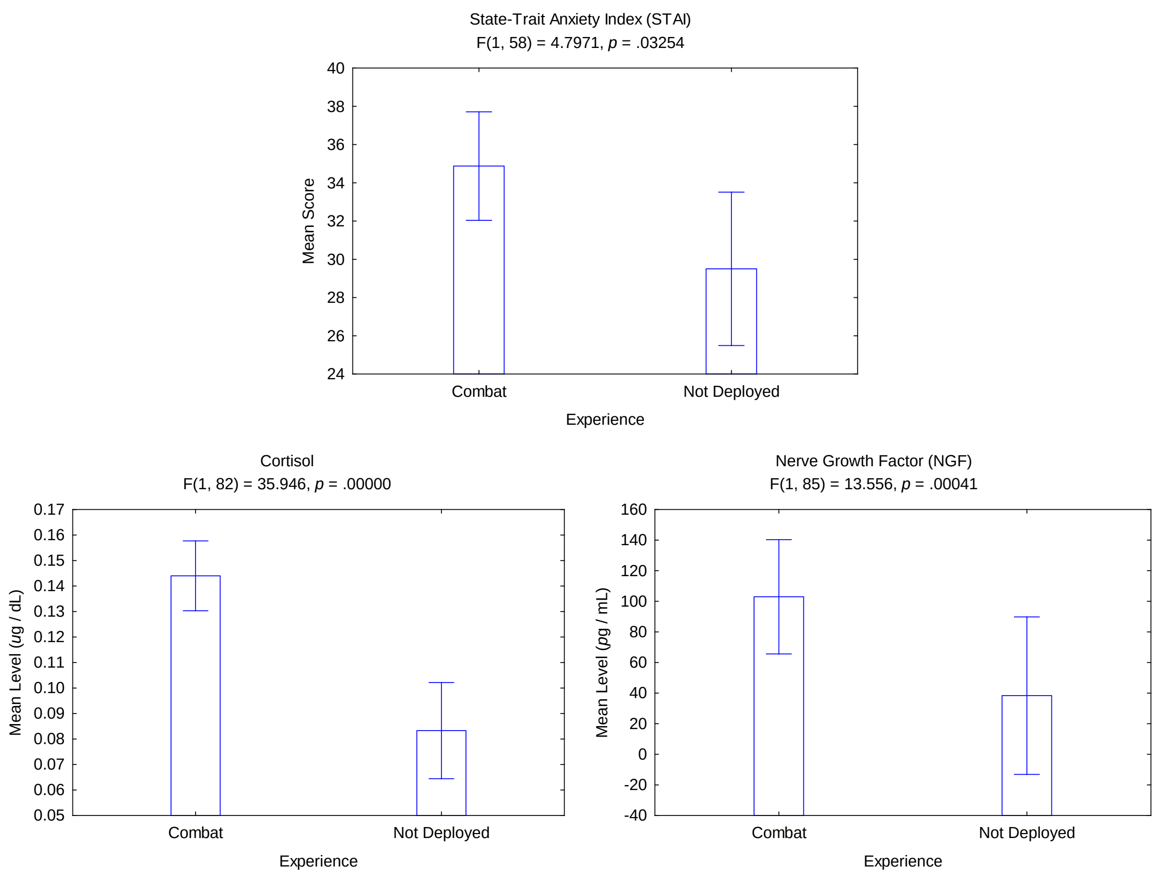

Figure 11. ANOVA Results: Sub-group Comparisons – Combat and Not Deployed.

Physiology data in these figures show that the stress response was higher for Marines with combat experience, when compared with Marines who had never been operationally deployed.

Combat versus Deployment

Comparisons were next made between Marines with combat experience (Combat) and Marines who had been operationally deployed, but who had not received a combat rating (Deployed). ANOVA results for this two-group test are shown in Figure 12.

Although self-report measures again failed to achieve statistical significance in this comparison, it is important to note that considerable variability was observed in the response data of the Deployed sub-group, which eliminated any potential effect. Only additional data collection can resolve this issue.

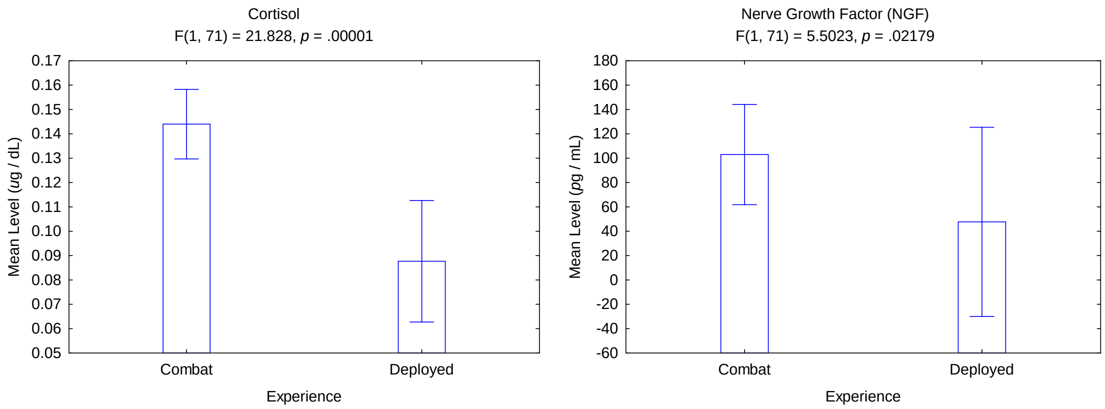

Figure 12. ANOVA Results: Sub-group Comparisons – Combat and Deployed.

As shown in the figure, physiological measures effectively distinguished between these two Marine sub-groups and, again, both measures reflected the same relationship—greater stress response for Marines with combat experience than for Marines with only deployment experience.

Deployment Only versus No Deployment

Finally, Marines who had been operationally deployed only (Deployed) were compared with Marines who had never been deployed (Not Deployed). No significant effects were found for these comparisons.

Summary

Sub-group tests yielded significant, interpretable effects for physiological measures only, with cortisol showing stronger patterns than NGF. Self-report measures failed to show any consistent effects, possibly due to the homogenous nature of the Marine cohort, at least on a psychological level. The distinction between the two most extreme samples—combat-experienced Marines and Marines who had no deployment experience—may be worth further exploration with these and other self-report measures, as this comparison may have operational and training utility for the military. This level of investigation appears to represent an analytical limit, however, at least for this experiment design and the number of subjects tested.

OBJECTIVE V – EVALUATING STRESS RECOVERY

Generation and exploration of the stress response among Marine and civilian communities was the primary focus of this experiment. The full study design, however, included a data collection point (*Recovery*) at a fixed time following the end of the last VBS2 task administration (see Figure 3 and Figure 4), after the participants had been told that the stress portion of the experiment had ended. The *Recovery* collection was included to confirm that individual stress levels had returned to a range equivalent to the initial *Baseline* (or *Training*) point, and to investigate possible group differences in the pattern of such recovery.

The time profiles for the major stress response measures are again shown in Figure 13, with the addition of the *Recovery* data point. As shown in the figure, all measures, including cortisol, reflect a downward trend from the preceding *Stress 2* measurement. Note that the Control group is not shown in these illustrations as the original experiment from which the control data were taken included a different procedure prior to the *Recovery* collection point, making a graphical comparison invalid.

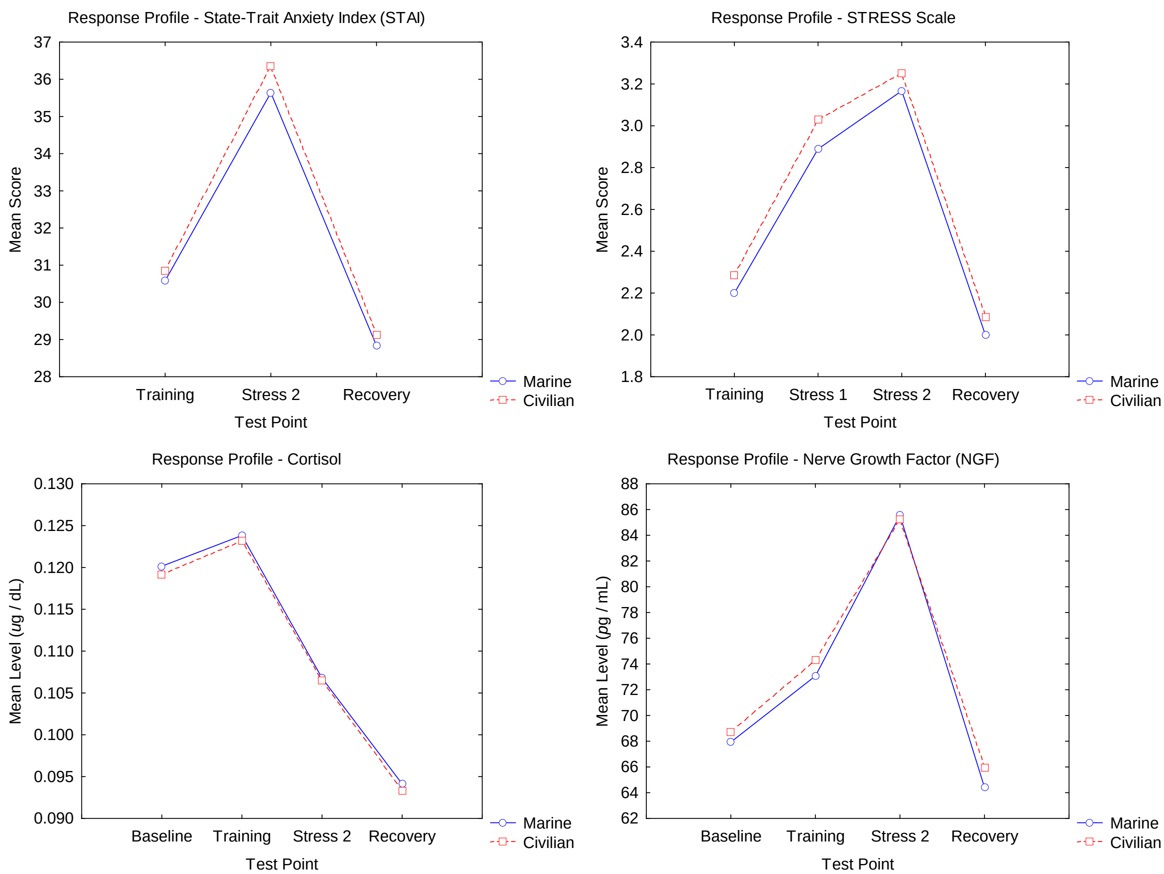

Figure 13. Time Profiles: Stress Response Measures Including Recovery.

With the exception of cortisol, which demonstrated an unusual dynamic throughout the study, these results behaved as expected. The cortisol pattern adds further evidence that the interpretation of physiology measures must address the dynamics of the generating mechanism for each substrate.

General Effects

One-way ANOVA tests for all groups (identical to the comparisons shown in Figure 6) revealed only a single significant result for the Recovery effect, NGF, as shown in Figure 14.

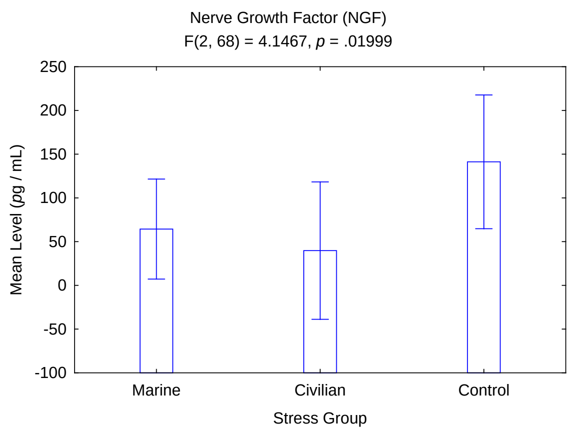

Figure 14. ANOVA Results for Recovery – All Group Comparison.

This unusual result, involving only NGF, did not warrant further analysis of pair-wise comparisons of *Recovery* among the primary Marine and civilian groups, and the general analysis was concluded at this point.

Experience Effects – Marine Sub-groups

Although general group effects were not encouraging, examination of recovery patterns *within* the Marine group was still of interest and might highlight additional community-unique effects. One-way ANOVA tests were therefore performed on all Marine sub-groups, with results as shown in Figure 15.

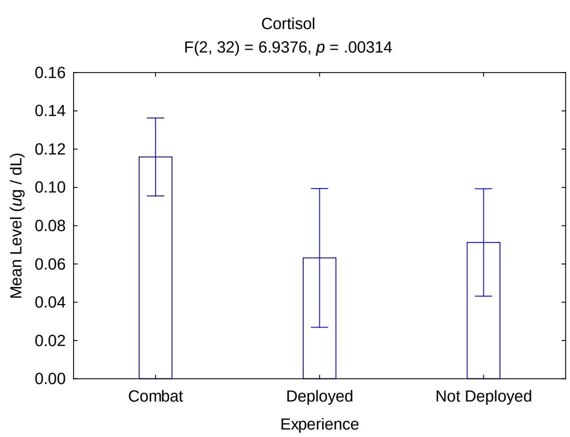

Figure 15. ANOVA Results for Recovery: Sub-group Comparisons – All Marines.

This result was also unusual, involving only cortisol as a significant stress measure. As shown in the figure, cortisol levels were higher for combat-experienced Marines at the *Recovery* point, compared to either of the other two Marine sub-groups, an effect that is confirmed with pair-wise comparisons shown in Figure 16.

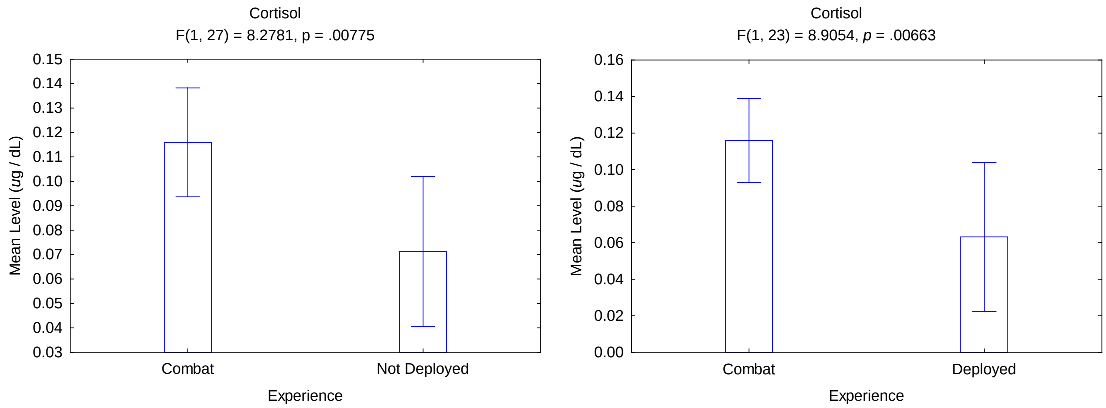

Figure 16. ANOVA Results for Recovery: Marine Sub-group Comparisons.

Summary

Although the analysis of recovery patterns yielded only limited results, involving a single stress measure—cortisol—the dynamics of stress recovery dynamics may have many contributing factors that would account for these modest effects, and further analysis is essential. Within these constraints, however, consistent physiological evidence was found to indicate that operational experience may impact the recovery from stress, as well as its onset and severity.

DISCUSSION

Collectively, the data from this experiment support the potential for detecting and characterizing stress response in both general (civilian) and specialized (Marine) groups, using both psychological and physiological methods. None of the measurement methods, however, was sufficiently sensitive or robust to demonstrate consistently significant or consistently interpretable outcomes. Self-report measures proved to be insightful when significant results were found, but response variability reduced the power of these measures in many cases where results would have been most valuable. Physiological measures achieved significance more than psychological self-report tools, but the patterns were often difficult to interpret or to relate to previous literature. No single tool or class of tools provided a general index of stress response.

GENERAL TESTS – MARINES AND CIVILIANS

The best overall results were obtained from comparisons of Marines, civilians who had completed the stress task, and a control group. As noted earlier, many ANOVA results showed intriguing patterns for psychological measures but failed to reach statistical significance because of large response variability. Although it is possible to obtain greater precision with larger data sets, and to resolve the significance of many effects, the need for large sample sizes reduces the value of a stress measure for operational field use (i.e., where only small groups may need to be evaluated).

The available psychological evidence from this series of tests showed that the stress response among Marines was lower than that of a Civilian Stress group but higher than that of a control group. This may be explained in terms of a <u>skill</u> difference: Because Infantry training likely provided the Marines with appropriate skills for the VBS2 stress task (compared to civilians), then it is reasonable to conclude that Marines brought a greater sense of control to the task, resulting in a lower stress response. This is, however, an example of appraisal; the physiological results for the same test series demonstrated higher stress responses for Marines than for the Civilian Stress group, which may indicate that conscious psychological appraisal and autonomic body mechanisms of stress reactivity represent different phenomena. Such a result would require the researcher to first generate a precise definition of *stress* in order to decide which measurement method—psychological or physiological—would provide the most appropriate information.

EVALUATING EXPERIENCE – MARINE SUB-GROUPS

Physiological measures showed far more consistent and interpretable patterns than self-report instruments when evaluating the (presumably) more homogenous sample of Marine participants. In fact, psychological measures proved of little value at this finer level of analysis, although this may be accounted for by a shared <u>ethos</u>, or value system, across the Marine group that influenced the psychological appraisal of task stress in this experiment.

In general, the examination of Marine sub-groups provided both psychological and physiological evidence for higher stress response among combat experienced Marines, compared with those who had not been deployed. To a smaller extent, physiological results also showed a similar relationship between combat-experienced Marines and those who had been deployed (but without a combat rating). These results more directly address the <u>understanding</u> aspect of the task relevance model introduced earlier, and therefore support the value of considering expertise in military communities as a predictive factor in stress response.

Other differences at this level of analysis are difficult to discriminate, which may mean that the sample sizes used for the experiment were too small to resolve them, that the measurement tools were not appropriate to or sensitive to such differences, or that such differences do not exist.

RECOVERY

Stress recovery was examined separately from the stress response, as the mechanisms underlying a return to baseline may be different than those which initially generate stress. Recovery patterns are important from both a military and scientific perspective. Comparison of Marine sub-groups provided interpretable results for cortisol, but the failure to observe these effects in other stress measures limits the promise of this line of inquiry, at least within the experiment design reported here.

CONCLUSIONS

This experiment was conducted to evaluate whether the stress responses of selected groups could be distinguished based on the properties of a stressor, to better understand the role of task content, as well as demand, on human reactions. Specifically, the study attempted to compare the effect of a difficult ground combat simulation task on the response of a group with (Marines) and without (civilians) a relevant skill set for the task. A secondary goal of the experiment was to compare the sensitivity and reliability of several stress measures from the research literature, to determine which approaches might be most effective as instruments for practical military field use.

In general, psychological stress responses were moderated and physiological responses were amplified for Marines, compared to civilians, and all responses were amplified for Marines with combat experience compared to other operational experience levels. These results could be interpreted as confirmation of a model of task relevance that included dimensions of task skill and task understanding as shaping factors.

Although each of the stress measures used for the experiment was effective in one or more levels of analysis, there was little consistency in their performance across the experiment. Two reasons that may account for these results include a distinction between (conscious) psychological appraisal and (autonomic) body responses to a stress event, and different generating mechanism for the physiological substrates used in the study. Until further work provides data to characterize and contrast the generative processes of these substrates as stress markers, the use of physiology for acute operational stress measurement will remain challenging.

REFERENCES

Aloe, L., Alleva, E., Bohm, A., and Levi-Montalcini, R. (1986). Aggressive behavior induces release of nerve growth factor from mouse salivary gland into the bloodstream. *Proceedings of the National Academy of Sciences, 83(16).*

Aloe, L., Alleva, E., and Fiore, M. (2002). Stress and nerve growth factor: Findings in animal models and humans. *Pharmacology Biochemistry and Behavior*, 73(1) 159-166.

Bland, J. M. (1995). Multiple significance tests: The Bonferroni method. *British Medical Journal*, 310, 170.

Bohemia Interactive (2010) http://www.bistudio.com/bohemia-interactive-simulations/virtual-battlespace-2_en.html.

Box, G.E., Hunter, W.G., and Hunter, J.S. (1978). Statistics for Experimenters. Hoboken, NJ: John Wiley and Sons.

Chiffer McKay, K.A., Buen, J.E, Bohan, K.J., and Maye, J.P. (2010). Determining the relationship of acute stress, anxiety, and salivary a-amylase level with performance of student nurse anesthesiologists during human-based anesthesia simulator training. *AANA Journal, 78(4).*

Driskell, J.E., and Salas, E. (1991). Overcoming the effects of stress on military performance: Human factors, training, and selection strategies. In R. Gal & D. Mangesldorf (Eds.), Handbook of Military Psychology. UK, Sussex: John Wiley and Sons.

Hancock, P. A., and Szalma, J. L. (2003). Operator stress and display design. *Ergonomics in Design*, 11, 13-18.

Kirschbaum, C., Klauer,T., Filipp, S.H., and Hellhammer, D.H. (1995). Sex-specific effects of social support on cortisol and subjective responses to acute psychological stress. *Psychosomatic Medicine, 57(1)*

Kirschbaum, C. and Hellhammer, D.H. (1999). Noise and stress – salivary cortisol as a non-invasive measure of allostatic load. *Noise Health* (1) 57-65

Lazarus, R.S., and Folkman, S. (1984). Stress, Appraisal and Coping. New York: Springer Publishing Company.

Leach, J. (2004). Why people “freeze” in an emergency: Temporal and cognitive constraints on survival responses. *Aviation, Space, and Environmental Medicine*, 75 (539-542).

Lieberman, H.R., Bathalon, G.P., Falco, C.M., Morgan III, C.A. Niro, P.J., and Tharion, W.J. (2005). The fog of war: Decrements in cognitive performance and mood associated with combat-like stress. *Aviation, Space, and Environmental Medicine*, 76(7), 7-14.

Murray, S.A., Ensign, W.E., and Yanagi, M.A. (2010). The effects of acute stress on cognitive performance: A pilot study. Technical Report 1995. Space and Naval Warfare Systems Center Pacific, San Diego, CA.

Noto, Y., Sato, T., Kudo, M., Kurata, K., and Hirota, K. (2005). The relationship between salivary biomarkers and state-trait anxiety inventory score under mental arithmetic stress: A pilot study. *Anesthesiology and Analgesia, 101* (1873-6).

Orosanau, J.M., and Backer, P. (1996). Stress and military performance. In J.E. Driskell & E. Salas (Eds.), Stress and Human Performance. Mahwah (NJ): Lawrence Erlbaum.

Ritter, F.E., Riefers, A.L., and Schoelles, M.J. (2005). Defining testable theories of pre-task appraisal stress.(Technical Report ACS 2005-2). Pennsylvania State University School of Information Sciences and Technology.

Salas, E., Driskell, J.E., and Hughes, J. (1996). Introduction: The study of stress and human performance. In J.E. Driskell & E. Salas (Eds.), Stress and Human Performance. Mahwah (NJ): Lawrence Erlbaum.

Spielberger, C.D. and Sydeman, S.J. (1994). State-Trait Anxiety Inventory and State-Trait Anger Expression Inventory. . In The Use of Psychological Testing for Treatment Planning and Outcome Assessment (Ch 13) 292-321. Hillsdale, NJ: Erlbaum Associates.

Steptoe, A., Hamer, M., and Chida, Y. (2007). The effects of acute psychological stress on circulating inflammatory factors in humans: A review and meta-analysis. *Brain, Behavior, and Immunity, 21(7).*

Stokes, A.F., Kemper, K.L., and Marsh, R. (1992). Time-stressed flight decision making: A study of expert and novice aviators (Technical Report ARL-93-1/INEL-93-1). Urbana-Champaign, IL: Aviation Research Laboratory, University of Illinois.

U.S. Army, (1998) Combat stress control in a theater of operations. C1, FM 8-51. Headquarters, Department of the Army Washington, DC, 30 January 1998.

Van Dongen, H.P., Baynard, M.D., Maislin, G., and Dinges, D.F. (2004). Systematic interindividual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability. *Sleep*, 27(3).

Vrijkotte, T.G.M., van Doornen, L.J.P., and de Geus, E.J.C. (2000). Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. *Hypertension*, 35(880).

Wang, Z., Rao, H., Wetmore, G.S., Furlan, P.M., Korczykowski, M., Dinges, D.F., and Detre, J.A. (2005). Perfusion functional MRI reveals cerebral blood flow pattern under psychological stress. *Proceedings of the national Academy of Sciences*, 102(49), 17804-17809.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-01-0188

1. REPORT DATE *(DD-MM-YYYY)*
December 2010

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Stress Response as a Function of Task Relevance

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Steven A. Murray, Matthew Yanagi, Wayne Ensign
SSC Pacific

Burcu Darst
Gray Sourcing

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

SSC Pacific
San Diego, CA 92152–5001

8. PERFORMING ORGANIZATION REPORT NUMBER

TR 1993

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
A laboratory study was conducted to evaluate the impact of task relevance—the relation between task characteristics and an individual's skill set—on the stress response of Marine Corps Infantrymen, and to evaluate the sensitivity of several stress measures to distinctions of task relevance. The stress task involved a difficult simulated ground combat mission using Virtual Battlespace 2 (VBS2), presented on a laptop computer. Stress response measures included two self-report questionnaires and two salivary hormones (cortisol and nerve growth factor) collected at several points in the testing session. Results confirmed that the VBS2 task was effective in generating a significant stress response. Differences in stress response were also found between Marines and civilians, and between sub-groups of Marines, that supported the concept of task relevance. Each stress measure proved effective in some, but not all, comparisons, which highlighted the caution required in applying and interpreting such measures, or in relying on any single tool for general stress evaluation. Results are discussed in terms of stress models and the potential of stress measurement tools for military field use.

15. SUBJECT TERMS
Mission Area: Research and Applied Sciences
Task relevance
Stress response measures

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U U		U	UU		Steven A. Murray
					19B. TELEPHONE NUMBER *(Include area code)* (619) 553–6350

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

INITIAL DISTRIBUTION

85300	Archive/Stock	(1)
71730	S. Murray	(1)

Defense Technical Information Center
Fort Belvoir, VA 22060–6218 (1)

SSC Pacific
San Diego, CA 92152-5001